

**MEPS HC-023:  
MEPS Panel I Longitudinal Weight File**

**July 2001**

**Agency for Healthcare Research and Quality  
Center for Cost and Financing Studies**

## Contents

A. Data Use Agreement .....	A-1
B. Background .....	B-1
1.0 Household Component .....	B-1
2.0 Medical Provider Component .....	B-2
3.0 Insurance Component .....	B-3
4.0 Nursing Home Component .....	B-3
5.0 Survey Management .....	B-4
C. Technical Information .....	C-1
1.0 Data File Contents .....	C-1
2.0 Variance Estimation .....	C-2

## **A. Data Use Agreement**

Individual identifiers have been removed from the micro-data contained in the files on this CD-ROM. Nevertheless, under sections 308 (d) and 903 (c) of the Public Health Service Act (42 U.S.C. 242m and 42 U.S.C. 299 a-1), data collected by the Agency for Healthcare Research and Quality (AHRQ) and /or the National Center for Health Statistics (NCHS) may not be used for any purpose other than for the purpose for which it was supplied; any effort to determine the identity of any reported cases, is prohibited by law.

Therefore in accordance with the above referenced Federal Statute, it is understood that:

1. No one is to use the data in this data set in any way except for statistical reporting and analysis; and
2. If the identity of any person or establishment should be discovered inadvertently, then (a) no use will be made of this knowledge, (b) The Director, Office of Management AHRQ will be advised of this incident, (c) the information that would identify any individual or establishment will be safeguarded or destroyed, as requested by AHRQ, and (d) no one else will be informed of the discovered identity.
3. No one will attempt to link this data set with individually identifiable records from any data sets other than the Medical Expenditure Panel survey or the National Health Interview Survey.

By using this data you signify your agreement to comply with the above stated statutorily based requirements with the knowledge that deliberately making a false statement in any matter within the jurisdiction of any department or agency of the Federal Government violates 18 U.S.C. 1001 and is punishable by a fine of up to \$10,000 or up to 5 years in prison.

The Agency for Healthcare Research and Quality requests that users cite AHRQ and the Medical Expenditure Panel Survey as the data source in any publications or research based upon these data.

## **B. Background**

This documentation describes one in a series of public use files from the Medical Expenditure Panel Survey (MEPS). The survey provides a new and extensive data set on the use of health services and health care in the United States.

The Medical Expenditure Panel Survey (MEPS) is conducted to provide nationally representative estimates of health care use, expenditures, sources of payment, and insurance coverage for the U.S. civilian non-institutionalized population. MEPS also includes a nationally representative survey of nursing homes and their residents. MEPS is cosponsored by the Agency for Healthcare Research and Quality (AHRQ) and the National Center for Health Statistics (NCHS).

MEPS comprises four component surveys: the Household Component (HC), the Medical Provider Component (MPC), the Insurance Component (IC), and the Nursing Home Component (NHC). The HC is the core survey, and it forms the basis for the MPC sample and part of the IC sample. The separate NHC sample supplements the other MEPS components. Together these surveys yield comprehensive data that provide national estimates of the level and distribution of health care use and expenditures, support health services research, and can be used to assess health care policy implications.

MEPS is the third in a series of national probability surveys conducted by AHRQ on the financing and use of medical care in the United States. The National Medical Care Expenditure Survey (NMCES, also known as NMES-1) was conducted in 1977, the National Medical Expenditure Survey (NMES-2) in 1987. Beginning in 1996, MEPS continues this series with design enhancements and efficiencies that provide a more current data resource to capture the changing dynamics of the health care delivery and insurance system.

The design efficiencies incorporated into MEPS are in accordance with the Department of Health and Human Services (DHHS) Survey Integration Plan of June 1995, which focused on consolidating DHHS surveys, achieving cost efficiencies, reducing respondent burden, and enhancing analytical capacities. To accommodate these goals, new MEPS design features include linkage with the National Health Interview Survey (NHIS), from which the sampling frame for the MEPS HC is drawn, and continuous longitudinal data collection for core survey components. The MEPS HC augments NHIS by selecting a sample of NHIS respondents, collecting additional data on their health care expenditures, and linking these data with additional information collected from the respondents' medical providers, employers, and insurance providers.

### **1.0 Household Component**

The MEPS HC, a nationally representative survey of the U.S. civilian non-institutionalized population, collects medical expenditure data at both the person and household levels. The HC collects detailed data on demographic characteristics, health conditions, health status, use of medical care services, charges and payments, access to care, satisfaction with care, health insurance coverage, income, and employment.

The HC uses an overlapping panel design in which data are collected through a preliminary contact followed by a series of five rounds of interviews over a 2 ½-year period. Using computer-assisted personal interviewing (CAPI) technology, data on medical expenditures and use for 2 calendar years are collected from each household. This series of data collection rounds is launched each subsequent year on a new sample of households to provide overlapping panels of survey data and, when combined with other ongoing panels, will provide continuous and current estimates of health care expenditures.

The sampling frame for the MEPS HC is drawn from respondents to NHIS, conducted by NCHS. NHIS provides a nationally representative sample of the U.S. civilian non-institutionalized population, with oversampling of Hispanics and blacks.

## **2.0 Medical Provider Component**

The MEPS MPC supplements and validates information on medical care events reported in the MEPS HC by contacting medical providers and pharmacies identified by household respondents. The MPC sample includes all hospitals, hospital physicians, home health agencies, and pharmacies reported in the HC. Also included in the MPC are all office-based physicians:

- Providing care for HC respondents receiving Medicaid.
- Associated with a 75-percent sample of HC households receiving care through an HMO (health maintenance organization) or managed care plan.
- Associated with a 25-percent sample of the remaining HC households.

Data are collected on medical and financial characteristics of medical and pharmacy events reported by HC respondents, including:

- Diagnoses coded according to ICD-9-CM (9th Revision, International Classification of Diseases) and DSM-IV (Fourth Edition, *Diagnostic and Statistical Manual of Mental Disorders*).
- Physician procedure codes classified by CPT-4 (Common Procedure Terminology, Version 4).
- Inpatient stay codes classified by DRGs (diagnosis-related groups).
- Prescriptions coded by national drug code (NDC), medication names, strength, and quantity dispensed.
- Charges, payments, and the reasons for any difference between charges and payments.

The MPC is conducted through telephone interviews and mailed survey materials.

### **3.0 Insurance Component**

The MEPS IC collects data on health insurance plans obtained through employers, unions, and other sources of private health insurance. Data obtained in the IC include the number and types of private insurance plans offered, benefits associated with these plans, premiums, contributions by employers and employees, eligibility requirements, and employer characteristics.

Establishments participating in the MEPS IC are selected through four sampling frames:

- A list of employers or other insurance providers identified by MEPS HC respondents who report having private health insurance at the Round 1 interview.
- A Bureau of the Census list frame of private sector business establishments.
- The Census of Governments from Bureau of the Census.
- An Internal Revenue Service list of the self-employed.

To provide an integrated picture of health insurance, data collected from the first sampling frame (employers and insurance providers) are linked back to data provided by the MEPS HC respondents. Data from the other three sampling frames are collected to provide annual national and State estimates of the supply of private health insurance available to American workers and to evaluate policy issues pertaining to health insurance.

The MEPS IC is an annual panel survey. Data are collected from the selected organizations through a prescreening telephone interview, a mailed questionnaire, and a telephone follow up for nonrespondents.

### **4.0 Nursing Home Component**

The 1996 MEPS NHC was a survey of nursing homes and persons residing in or admitted to nursing homes at any time during calendar year 1996. The NHC gathered information on the demographic characteristics, residence history, health and functional status, use of services, use of prescription medications, and health care expenditures of nursing home residents. Nursing home administrators and designated staff also provided information on facility size, ownership, certification status, services provided, revenues and expenses, and other facility characteristics. Data on the income, assets, family relationships, and care-giving services for sampled nursing home residents were obtained from next-of-kin or other knowledgeable persons in the community.

The 1996 MEPS NHC sample was selected using a two-stage stratified probability design. In the first stage, facilities were selected; in the second stage, facility residents were sampled, selecting both persons in residence on January 1, 1996, and those admitted during the period January 1 through December 31.

The sample frame for facilities was derived from the National Health Provider Inventory, which is updated periodically by NCHS. The MEPS NHC data were collected in person in three rounds of data collection over a 1 ½-year period using the CAPI system. Community data were collected by telephone using computer-assisted telephone interviewing (CATI) technology. At the end of three rounds of data collection, the sample consists of approximately 815 responding facilities, 3,100 residents in the facility on January 1, and 2,200 eligible residents admitted during 1996.

## **5.0 Survey Management**

MEPS data are collected under the authority of the Public Health Service Act. They are edited and published in accordance with the confidentiality provisions of this act and the Privacy Act. NCHS provides consultation and technical assistance.

As soon as data collection and editing are completed, the MEPS survey data are released to the public in staged releases of summary reports and microdata files. Summary reports are released as printed documents and electronic files. Microdata files are released on CD-ROM and/or as electronic files. A catalog of all MEPS products released to date is provided in Section F of this document.

Printed documents and CD-ROMs are available through the AHRQ Publications Clearinghouse. Write or call:

AHRQ Publications Clearinghouse  
Attn: (publication number)  
P.O. Box 8547  
Silver Spring, MD 20907  
800/358-9295  
410/381-3150 (callers outside the United States only)  
888/586-6340 (toll-free TDD service; hearing impaired only)

Be sure to specify the AHRQ number of the document or CD-ROM you are requesting. Selected electronic files are available from the Internet on the MEPS home page:  
<http://www.meps.AHRQ.gov/>.

Additional information on MEPS is available from the MEPS project manager or the MEPS public use data manager at the Center for Cost and Financing Studies, Agency for Healthcare Research and Quality.

## C. Technical Information

### Overview

This documentation describes a two-year Longitudinal File derived from the respondents to the Medical Expenditure Panel Survey (MEPS) Panel I 1996 sample. The MEPS, a nationally representative survey of the U.S. civilian non-institutionalized population, uses a panel design in which data were collected through a preliminary contact followed by a series of five rounds of interviews to collect data for a 2 year period. The persons on the dataset represent those who were in this population for all or part of the 1996-1997 period. To obtain analytic variables, the records on this file need to be linked to the 1996 and 1997 MEPS public use data sets by the sample person identifier (DUPERSID).

### 1.0 Data File Contents

This file contains a total of 19,859 persons, of whom 19,289 are available for longitudinal analysis over a 2-year period. There are also 237 people on the file who provided data only in Year 1. These are people who died, went into a nursing home, or became ineligible for the survey for some other reason, such as they entered into the military or left the country. In addition, there are 333 people of the file who provided data only in Year 2. These are newborns or those who came into a selected household for the first time in 1997, such as those who moved into a sample household from a nursing home or other institution. The additional sample that exists for only 1 year is provided on the file to facilitate analyses that cover the experience of the U.S. civilian non-institutionalized population over 1996 and 1997.

This file provides a weight variable (LONGWGT) that, when applied to the persons who participated in both 1996 and 1997, will enable the user to make national estimates of person-level changes in selected variables [e.g., health insurance, health status, utilization and expenditures]. In addition, LONGWGT can also be used to develop cross-sectional type estimates for the civilian noninstitutionalized population in each year based on only the panel 1 sample.

These estimates are robust and similar to those constructed using the standard 1996 and 1997 weights (WTDPER96 and WTDPER97) in the MEPS public use files. **However, if the purpose of the analysis is only to produce estimates for these survey years independently, it is preferable to use the existing Public Use Files (HC012 for 1996 and HC020 for 1997).** The 1997 PUF file has a larger sample size and will therefore produce estimates with smaller variances.

The estimate of total health care expenditures for 1997 using the longitudinal weight is \$529.5 billion. Using the 1997 Public Use File (HC020), the estimate of total health care expenditures is \$553.2 billion. While these estimates are not statistically significantly different, an overall adjustment could be made to improve the alignment across these estimates. To adjust mean or total expenditure estimates derived from this longitudinal file in order to replicate the overall estimates derived from the 1997 HC020 file within population subgroups (c) or for the overall population, it will be necessary to develop adjustment factors,  $A(c)$ , which are defined as the

ratio of the weighted estimate of health care expenditures derived from HC020 over the weighted health care expenditure estimate obtained from this file for subgroup c. For example, to derive a mean expenditure estimate  $\bar{Y}_c$  that is adjusted in this manner for subgroup c (e.g, for age group 65+), use the following method:

$$A(c) = ( \sum_{i \in c} W_{2i} Y_i ) / ( \sum_{i \in c} W_{1i} Y_i ), \text{ and}$$

$$\bar{Y}_c = ( \sum_{i \in c} A(c) W_{1i} Y_i ) / ( \sum_{i \in c} W_{1i} )$$

where

$Y_i$  is the expenditure variable of interest for individual i,  
 $W_{1i}$  is the longitudinal weight for individual i,  
 $W_{2i}$  is the person weight from HC020 for individual i, and  
the sum is across all sample participants in group c.

The following table contains a summary of cases to include, sample sizes, and population estimates (i.e. sum of LONGWGT) for the 3 different time periods.

Population of Interest	Cases to Include	Sample Size	Population Estimate
1996-1997	YEARIND=1	19,289	266,895,227
1996 Only	YEARIND=1 or 2	19,526	270,220,553
1997 Only	YEARIND=1 or 3	19,622	271,278,585

## 2.0 Variance Estimation

To obtain estimates of variability (such as the standard error of sample estimates or corresponding confidence intervals) for estimates based on MEPS survey data, one needs to take into account the complex sample design of MEPS. The variables needed to implement a Taylor series estimation approach are included on the Longitudinal File. They are VARSTR and VARPSU. These variables can be used for producing 1996 or 1997 estimates.